Landes Bioscience

www.landesbioscience.com

Supplemental Material to:

Timo Quante, Benjamin Otto, Marie Brázdová,
Iva Kejnovská, Wolfgang Deppert and Genrich Tolstonog

Mutant p53 is a transcriptional co-factor that binds to G-rich regulatory regions of active genes and generates transcriptional plasticity

2012; 11(17)

http://dx.doi.org/10.4161/cc.21646

http://www.landesbioscience.com/journals/cc/article/21646

SUPPLEMENTAL DATA

Mutant p53 is a transcriptional co-factor that binds to G-rich regulatory regions of active genes and generates transcriptional plasticity

Timo Quante, Benjamin Otto, Marie Brázdová, Iva Kejnovská, Wolfgang Deppert and Genrich V. Tolstonog

SUPPLEMENTAL METHODS

siRNA transfection

U251 cells were seeded in 6-wells or 10-cm dishes at a density of 0.5×10^5 or 2.95×10^5 cells per well/dish 24 h before transfection. For each 6-well 10 µl 20 µM siRNA-duplex (Table S5) were mixed with 175 µl OptiMEM (Invitrogen). In another tube 11 µl of OptiMEM were mixed with 4 µl Oligofectamine (Invitrogen) by pipetting, kept at RT for 10 min and then added to the diluted siRNA, again mixed by pipetting and kept at RT for another 20 min. For a 10-cm dish 60 µl 20 µM siRNA-duplex were mixed with 1050 µl OptiMEM and 66 µl OptiMEM were mixed with 24 µl Oligofectamine. A mock control with water instead of siRNA was included. The cells were washed once with warm PBS, 1.3 ml or 7.67 ml (6-well or 10-cm dish) of DMEM with 10% FCS was added and 200 µl or 1200 µl of the prepared siRNA/Oligofectamine solution was added drop-wise followed by gentle shaking of the plate. The cells were incubated for 48 hours before RNA and protein were extracted for Real-Time PCR and western blotting. For later time points and ChIP-assays the cells were splitted 1:2 to 1:3 48 h after transfection and reseeded to 6-wells or 15-cm dishes.

Neurosphere culture

Cells were harvested by trypsination and counted. 2.5×10^5 cells were resuspended in Neurobasal A medium (Gibco) supplemented with B27 supplement (Gibco), 0.1 % BSA, 10 ng/ml bFGF (P.J.K. GmbH) and 20 ng/ml EGF (P.J.K. GmbH) and seeded in Ultra Low Cell Culture Flasks (Cornig). For passaging neurospheres were collected by centrifugation at 300 x g for 5 min. Before re-seeding single cell suspensions were prepared by incubating the pellet for 2 times 5 min at 37°C in 500 µl trypsin/EDTA with extensive resuspenion by pipetting and subsequent filtration through a 30 µm filter.

Agar cloning

Seeding and culture of cells in soft agar was performed according to standard procedures. After 2-3 weeks in the CO_2 incubator, colonies were counted.

Immunoblotting

Proteins were extracted after washing the cells once with warm PBS by direct lysis in 200 µl (per 6-well) of Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 5% ß-mercaptoethanol, 0.1% bromphenol blue). Proteins were separated on SDS-polyacrylamide gels, electroblotted onto PVDF membranes (Millipore) and detected using anti-p53, anti-HSC70 and anti-alpha-tubulin antibodies (Table S5). HRP-conjugated donkey anti-goat, donkey anti-rabbit, donkey anti-mouse and donkey anti-rat antibodies were obtained from Dianova. Proteins were visualized using SuperSignal West Dura Chemiluminescent Substrate (Thermo Scientific).

Chromatin immunoprecipitation (ChIP)

Cells were cultured in 15-cm diameter tissue culture dishes. At 80% confluence, cells were washed with warm PBS (pH 7.4) and were either treated directly with DMEM supplemented with 1%

formaldehyde (Sigma-Aldrich) for 10 min at room temperature or subjected to a nuclear extraction procedure (see Supplementary Methods) to remove the cytosolic fraction before crosslinking. Crosslinking was stopped by adding 125 mM glycin (final concentration). Treated cells were washed twice with cold PBS, scraped in PBS and collected by centrifugation at 1350 x g at 4°C for 5 min. Cell pellets from two to five 15-cm dishes were pooled and resuspended in 5 ml lysis buffer 1 (140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100 in 50 mM HEPES-KOH, pH 7.5 supplemented with protease inhibitors (5 µg/ml leupeptin (Sigma-Aldrich), 1 µg/ml pepstatin (Sigma-Aldrich), 1% Trasylol (Bayer), 0.1 mg/ml Pefabloc SC (Biomol GmbH)) and incubated for 10 min on ice with gentle shaking. After centrifugation at 1350 x g at 4°C for 5 min, the supernatant was discarded and the pellet was resuspended in 5 ml lysis buffer 2 (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA in 10 mM Tris-HCl, pH 8.0 supplemented with protease inhibitors) and left shaking at room temperature for 10 min. Cells were pelleted by centrifugation at 1350 x *g* at 4°C for 5 min, the supernatant was discarded and the pellet was resuspended in 2-3 ml lysis buffer 3 (100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine in 10 mM Tris-HCl, pH 8.0 supplemented with protease inhibitors) or ChIP-SDS-lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% (w/v) SDS) and sonicated in 15 ml Falcon tubes for 15 minutes in a Bioruptor UCD-200 (Diagenode) with 30 second on/off cycles. After sonication, when lysis buffer 3 had been used Triton X-100 was added to the samples to a final concentration of 1% and debris was eliminated by centrifugation for 20 minutes at 20,000 x *g* at 4°C. Samples in ChIP-SDS-lysis buffer were centrifuged without addition of Triton X-100. 50 µl of the lysate was collected as input and after addition of 150 µl elution buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0) the crosslinks were reversed by incubation at 65°C overnight. 200 µl TE buffer and 8 µl 10 mg/ml RNAse A were added followed by 2 hours at 37°C. To remove proteins 300 mM $CaCl_2$ in 10 mM Tris pH 8.0 and 4 µl 20 mg/ml proteinase K (Roche) were added and samples were incubated for at least 30 min at 50°C. DNA was purified by phenol-chloroform-isoamylalcohol (25:24:1; Biomol GmbH) extraction with subsequent precipitation with ethanol in the presence of NaCl (200 mM final concentration) and 20 µg glycogen (Roche), washed with 70% ethanol, air-dried and dissolved in water. DNA size was estimated by agarose gel electrophoresis.

For immunoprecipitation (IP), anti-histone H3 lysine 4 trimethylation, anti-histone H3 lysine 9 acetylation, anti-RNA-Polymerase II, anti-RNA-Polymerase II CTD repeat YSPTSPS phospho S2, anti-RNA-Polymerase II CTD repeat YSPTSPS phospho S5, anti-p53, anti-SP1, and anti-ETS1 antibodies (Table S5) and protein A or anti-mouse IgG Dynabeads (Invitrogen) were used. Beads and antibodies were pre-incubated in blocking buffer (0.5% BSA in PBS) overnight at 4°C with shaking and washed three times in blocking buffer prior to IP. 100-200 µl sonicated lysate in lysis buffer 3 were used per IP and brought to 1 ml reaction volume by addition of IP buffer (100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine, 1% Triton X-100 in 10 mM Tris-HCl, pH 8.0). In case of samples in ChIP-SDS-lysis buffer 100-200 µl sonicated lysate were diluted 1:10 in ChIP dilution buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1 % (v/v) Triton X-100). The pre-incubated antibody/beads were added and samples were incubated overnight at 4°C with agitation. The next day beads were washed three times with IP buffer, once in high salt wash buffer (500 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine, 1% Triton X-100 in 10 mM Tris-HCl, pH 8.0), once in RIPA buffer (500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-

Deoxycholate in 50 mM HEPES-KOH, pH 7.6) and once in 50 mM NaCl in TE buffer. Protein-DNA complexes were eluted in 210 µl elution buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0) for 15 minutes at 65°C and supernatants were collected. Crosslink reversal, RNAse A and proteinase K treatment and DNA purification were performed as described for the input samples.

Nuclear extraction for ChIP

Cell lysates were prepared from 3-5 15-cm cell culture dishes at 80 % confluency. Cells were washed once with warm PBS, trypsinized and pelleted at 300 x g for 5 min at 4°C. The pellet was washed once in cold PBS and resuspended in 2 ml of cold nuclei buffer 1 (0.3 M sucrose in 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EGTA, 15 mM Tris-HCl pH 7.5, supplemented with 0.5 mM DTT and protease inhibitors (10 µg/ml leupeptin (Sigma-Aldrich), 1 µg/ml pepstatin (Sigma-Aldrich), 1% Trasylol (Bayer), 0.1 mg/ml Pefabloc SC (Biomol GmbH)). After addition of 2 ml of cold nuclei buffer 2 (0.3 M sucrose in 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EGTA, 0.4 % NP-40, 15 mM Tris-HCl pH 7.5, supplemented with 0.5 mM DTT and protease inhibitors) and careful mixing cells were incubated for 10 min on ice. In the meantime two 15 ml Falcon tubes with 8 ml of cold nuclei buffer 3 (1.2 M sucrose in 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EGTA, 15 mM Tris-HCl pH 7.5, supplemented with 0.5 mM DTT and protease inhibitors) were prepared. 2 ml of the lysate was layered carefully on top of nuclei buffer 3 in the Falcon tubes and nuclei were collected by centrifugation at 4°C with 10000 x g for 20 min. The supernatant was removed carefully, the nuclei were washed once with cold PBS and subjected to crosslinking.

Bioinformatic analysis of ChIP-chip data

Scaled log2 ratios (Cy5-IP/Cy3-input) were calculated using the NimbleScan software (version 2.6.0). Preprocessing of Nimblegen GFF files (quantile normalization, mean signal summarization) and binding peak identification were performed using the Carpet Tools[1] provided by the IFOM-IEO Campus GALAXY platform compilation (http://host13.bioinfo3.ifom-ieo-campus.it:8080/). For peak detection in mutp53, SP1 and ETS1 ChIP-chip experiments the "p-value" analysis with the following settings was used:-log p-value cutoff between 1.3 and 4 (3 for histone modifications and Pol II S2-P/S5-P), a percentile value of 0.9 (0.8 for histone modifications and Pol II S2-P/S5-P), a minimal number of 4 probes, a max distance of 200 bp between two probes and a minimum distance of 200 bp between two peaks as well as a window length of 1 kb (Table S2 and Table S4).

The Table Browser function (http://genome.ucsc.edu) was used to retrieve genomic coordinates of CpG islands (cpgIslandExt track) and DNAseI-HS regions (wgEncodeRegDnaseClustered track) from the GRCh37 (hg19) human genome assembly. For analysis of intersections with binding regions identified in ChIP-chip studies we used the top 20% (items in cluster) regions from the DNAseI-HS track. The coordinates of G-quadruplex forming repeats were retrieved from the non-B DB[2] (http://nonb.abcc.ncifcrf.gov/), a database that contains genomic coordinates for predicted non-B DNA forming repeats. Coordinates of SP1 and ETS1 binding peaks were retrieved from the ENC TF Binding Supertrack which contains ChIP-Seq data from the ENCODE project (http://genome.ucsc.edu/ENCODE/index.html). This track contains coordinates of binding peaks computed from multiple experiments (different cell lines and transcription-factor targeting antibodies). Individual experiments containing SP1 (five cell lines, two replicates each) and ETS1 (three cell lines, two replicates each) binding sites identified by ChIP-seq were downloaded from the ENCODE/HAIB track[3]. The Ensembl transcript coordinates (human genome assembly GRCh37.p3) were retrieved from Biomart database using the MartView option (http://www.biomart.org/martservice.html). We used a list of Ensembl transcript coordinates that was restricted to 'regular' chromosomes (excludes haplotype chromosomes, unplaced contigs and mitochondrial chromosome) and to 'processed_transcript' and 'protein_coding' transcript biotypes. The annotation of the identified peaks was performed using GALAXY tools[4, 5]. Scripting and data visualization was done with R statistical platform (ver. 2.12.2) and Integrative Genomics Viewer (IGV) software (version 2.0)[6].

To analyze the overlay between two binding regions (e.g. mutp53 and SP1) or between binding regions and CpG islands and DNAseI-HS regions an intersection score was calculated according to

$$score = \frac{r - n * (R/N)}{\sqrt{n * (R/N) * (1 - (R/N)) * \left(1 - \frac{(n-1)}{(N-1)}\right)}}$$

where *N* is the median size of mapped transcripts (background size for peaks), *R* the first peak size, *n* the second peak size and *r* the size of the overlapping region.

Histone modifications (H3K4me3, H3K9ac) and RNA Pol II S2-P/S5-P binding frequencies near mutp53-BRs were determined by identification of corresponding peaks within a window of ±4 kb around the mutp53-BR center positions and the occurrence frequency was plotted dependent on the distance to the mutp53-BR center position. For estimation of the overlap ratios between mutp53-BRs and histone modifications (H3K4me3, H3K9ac) as well as RNA Pol II S2-P/S5-P binding peaks mutp53-BRs were restricted to those with a maximum distance of ±2 kb to the TSS region. Afterwards all modification peaks with a maximum distance of 200 bp to the flanks of these mutp53-BRs were selected. Finally the ratio, the overlap size and the mutp53-BR size was calculated (Table S6).

MEME-ChIP web-tool (http://meme.nbcr.net) was used to identify overrepresented nucleotide motifs[7]. Mutp53 binding regions were analyzed using 30 nt as maximum motif width, for ETS1 and SP1 binding regions the motif width was reduced to 10 nt.

EpiGRAPH web-tool[8] (http://epigraph.mpi-inf.mpg.de) was used to compare mutp53-binding regions with a background set. The random regions used as background set were generated in a way that the median peak size and number of peaks corresponded to the values observed for mutp53-BR retrieved from the ChIP-chip experiments. The calculation was performed in four steps: determination of potential region seed locations, selection of random region seeds, calculation of the raw binding peak borders and final correction of the peak borders to only generate random regions that are present on the array used for ChIP-chip. Using the R random seed generator a number of seeds depending on the number of peaks observed in the experiments (mutp53 = 517, ETS1 = 434, SP1 = 374) were then selected. The median binding peak size was determined for each transcription factor (mutp53 = 993, ETS1 = 748, SP1 = 729) and the selected seeds were extended to peaks of the according median size centered on the seed base pair. The peak borders were finally adjusted to start and end sites of the flanking oligonucleotides mapped on the array.

Quantitative Real Time PCR

Total RNA was purified using peqGOLD TriFast (Peqlab) and reverse transcribed with the High Capacity RT kit (Applied Biosystems) according to the manufacturer's protocol. PCR was performed using the Power SYBR Green PCR Master Mix (Applied Biosystems) in a standard program (10 min 95°C; 15 sec 95°C, 1 min 60°C; 40 cycles) running in an ABI 7900HT Fast thermal cycler (Applied Biosystems). PCR reactions for each sample were repeated in triplicates. The integrity of the amplified products was confirmed by melting-curve analysis. PCR primers (Table S5) were designed using the Primer3 web tool (http://frodo.wi.mit.edu/primer3/PCR). PCR efficiency was measured for each primer pair using serial dilutions of cDNA. The housekeeping gene *GAPDH* was used as endogenous control. Relative quantitation of transcript levels was done based on the 2^{-ddCt} method.

Immunofluorescence microscopy

Cells were grown on glass coverslips to 80-90 % confluence, washed once with PBS and fixed in 4 % paraformaldehyde in PBS for 10 min at room temperature (RT). Permeabilization was performed in 1 % Triton X-100 in PBS for 10 min at RT. Subsequently the cells were washed three times with PBS and blocked with 0.5 % BSA in PBS for 30 min at RT. To analyze the chromatin bound fraction, cells were subjected to an extraction procedure prior to fixation. Coverslips were washed once each in cold PBS and cold extraction buffer I (10 mM PIPES pH 6.8, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 1 mM EGTA, 0.5 % Triton X-100) and incubated in extraction buffer I for 5 min on ice. Cells were then washed once in extraction buffer II (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 1% NP-40, 0.5 % sodium deoxycholate) and incubated for 5 min on ice in extraction buffer II. After one wash in PBS fixation and blocking were performed as described above. After blocking cells were co-stained with rabbit polyclonal anti-p53 (sc-6243, FL-393, Santa Cruz Biotechnology) or mouse monoclonal anti-p53 (DO1) and rabbit polyclonal anti-SP1 (#07-645, Upstate/Millipore) or mouse monclonal anti-ETS1 (sc-55581, Santa Cruz Biotechnology). Donkey anti-rabbit and anti-mouse fluorochrome-coupled antibodies were obtained from Molecular Probes. Nuclei were stained with DAPI. The images were captured on a Leica DMI6000 B microscope equipped with a DFC350 FX camera.

Co-immunoprecipitation

Cell lysates were prepared from 3-6 15-cm cell culture dishes at 80 % confluency. Cells were washed once with warm PBS, scraped in 10 ml of cold PBS and pelleted at 300 x g for 5 min at 4°C. The pellet was resuspended in 1 ml of cold Co-IP Nuclei Buffer (10 mM Tris-HCl pH 7.9, 10 mM KCl, 15 mM $MgCl_2$, supplemented with protease inhibitors (5 µg/ml leupeptin (Sigma-Aldrich), 1 µg/ml pepstatin (Sigma-Aldrich), 1% Trasylol (Bayer), 0.1 mg/ml Pefabloc SC (Biomol GmbH)) and incubated for 10 min on ice with occasional mixing. Nuclei were collected by centrifugation at 4°C with 1000 x g for 2 min and according to the number of IPs to be done (100 µl lysate per IP, 50 µl input) the pellet was resuspended in 250-650 µl Co-IP lysis buffer (50 mM HEPES-KOH pH 7.5, 300 mM NaCl, 0.5 % (v/v) NP-40, supplemented with protease inhibitors and 3-6 µl 25 U/µl Benzonase). Lysis was performed for 30 min on ice with occasional mixing. Afterwards samples were centrifuged for 5 min at 4°C and 10,000 x g and supernatants were transferred to a new tube to remove remaining insoluble debris. To prepare magnetic beads with coupled antibodies for IP 40 µl of the appropriate beads (Dynabeads Protein A or anti mouse IgG, Invitrogen) were washed two times in 0.5 % BSA in PBS, resuspended in 250 µl 0.5% BSA in PBS and incubated at 4°C over night with agitation with 4 µg of the appropriate antibody (anti-p53, anti-SP1, anti-ETS1, rabbit IgG, and mouse IgG; Table S5). The next day the beads were washed two times in 0.5 % BSA/PBS and resuspended in 40 µl 0.5 % BSA/PBS. In case of the ETS1/p53 Co-IP the antibodies were coupled to the beads with BS3 crosslinking reagent (Thermo Fisher Scientific). For this purpose after incubation with the antibody the beads were washed two times in conjugation buffer (20 mM sodium phosphate, 150 mM NaCl, pH 8.0), resuspended in 250 µl 5 mM BS3 and incubated at room temperature for 30 min with agitation. The crosslinking reaction was quenched by adding 12.5 µl 1 M Tris-HCl pH 7.5 and the beads were washed three times in Co-IP dilution buffer (50 mM HEPES-KOH pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5 % (v/v) NP-40 supplemented with protease inhibitors. Per IP 100 µl of the lysate were diluted with 900 µl Co-IP dilution buffer. The samples were incubated overnight at 4°C with 40 µl magnetic beads that had been pre-incubated with the respective antibody. The next day the beads were washed five times with Co-IP dilution buffer, resuspended in 50 µl Co-IP elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% (w/v) SDS) and incubated for 15 min at 65°C and 900 rpm. Beads were collected on a magnet and the supernatant was transferred to a new reaction tube, 10 µl Laemmli buffer were added and the samples were boiled for 5 min at 95°C. 25 µl IP samples and 5 µl input were separated by SDS-PAGE on 12 % gels, blotted on nitrocellulose membranes and the respective proteins were detected in Western Blot.

DNA binding assays

Sequences of oligonucleotides used in this study (Table S5) were synthesized by VWS (Austria). G-quadruplexes were formed as described[9]. In brief, the oligonucleotides were radiolabeled with gamma-^{32}P dATP and T4 polynucleotide kinase (NEB) and single-stranded (ss) or double-stranded (ds) DNA was purified by mini Quick Spin DNA columns according to a standard protocol (Roche). P53CON and *MYC*-Py52/Pu52 duplex DNAs were prepared before purification by denaturating the oligonucleotides for 2 min at 98°C in T4 PNK buffer, than slowly cooling them down to room temperature. Radiolabeled *MYC*-Pu52, *TERT*-Pu61 were preassembled into a G-quadruplex via incubation in 10 mM Tris-HCl (pH 7.4) at 90°C for 5 min, then in 50 mM KCl, 10 mM Tris-HCl (pH 7.4) for 12 hours at 20°C. Recombinant wild-type and mutant (G245S, R248W, and R273H) human full-length p53 protein was expressed in *E.coli* strain C41(DE3)[10] using a two-step induction at 18°C to limit protein aggregation. Recombinant proteins were purified as described[11, 12]. The purity and integrity of recombinant proteins was analyzed by SDS-PAGE and western blotting using anti-p53 N-terminal and C-terminal specific antibodies (not shown). ^{32}P-radiolabeled oligonucleotide probe (2 pmol) was mixed with p53 proteins and incubated in binding buffer (5 mM Tris-HCl, pH 7.4, 0.01% Triton X-100, 0.5 mM EDTA, 50 mM or 100mM KCl, 1mM DTT, 5ng/µl BSA) in the presence of competitor DNA (20 ng SmaI-linearized pBluescript vector DNA) for 30 min at 25°C to reach equilibrium. Samples were loaded onto a 6% polyacrylamide gel containing 0.5x TBE buffer with 50 mM KCl. After 1 h of electrophoresis (at 6 V/cm), gels were dried and signals detected by autoradiography using Phosphorimager Storm (Typhoon FLA9000, GE). For nonradioactive EMSA 20 ng of the respective oligonucleotides were detected by SybrGold staining (Invitrogen). Except for the presence of competitor DNA all other conditions were the same as described above.

CD Spectroscopy

CD measurements were done on a Jobin-Yvon CD6 dichrograph in 1-cm path-length quartz Hellma cells placed in a thermostatted cell holder at 0°C; scan rate was 0.5 nm per second. The CD signal was expressed as the difference in the molar absorption of the right- and left-handed circularly polarized light. The molarity was related to strands. Precise DNA concentrations (0.5 μM *MYC*-Pu52, 0.9 μM *TERT*-Pu61) were determined on the basis of UV absorption at 260 nm of the sample measured in 5 mM Tris-HCl buffer (pH 7.4), 0.5 mM EDTA at 90°C, using molar extinction coefficients calculated as described [13]. UV absorption spectra were measured on a UNICAM 5625 UV/VIS spectrometer (Cambridge, UK). Experimental conditions were changed directly in the cells by adding solution (25 mM HEPES, pH 7.6, 200 mM KCl, 10% glycerol, 1 mM DTT) with or without the protein and the final DNA strand concentration was corrected according to the increase in volume.

SUPPLEMENTAL FIGURES

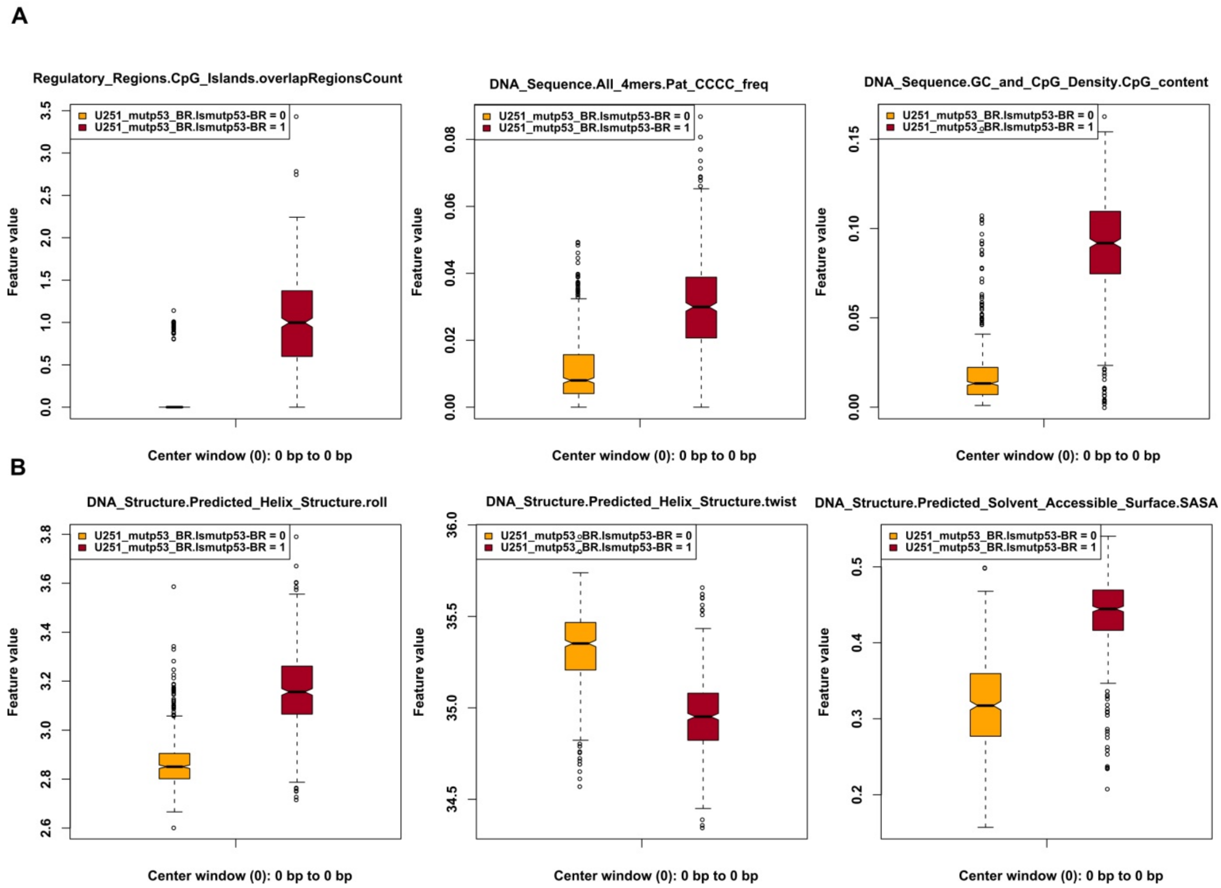

Figure S1 A, B. EpiGRAPH-generated boxplot diagrams demonstrating differences between mutp53-bound (red boxes) and random (yellow boxes) regions. 517 mutp53-bound regions (at -log p = 3) were compared with an equally sized random set of regions on the tiling array with respect to their overlap with CpG-islands, the presence of ‘CCCC’ pattern and the density of CpG dinucleotides (A). In (B) mutp53-bound and random regions were compared according to predicted parameters of DNA structure (roll, twist, and solvent accessible surface).

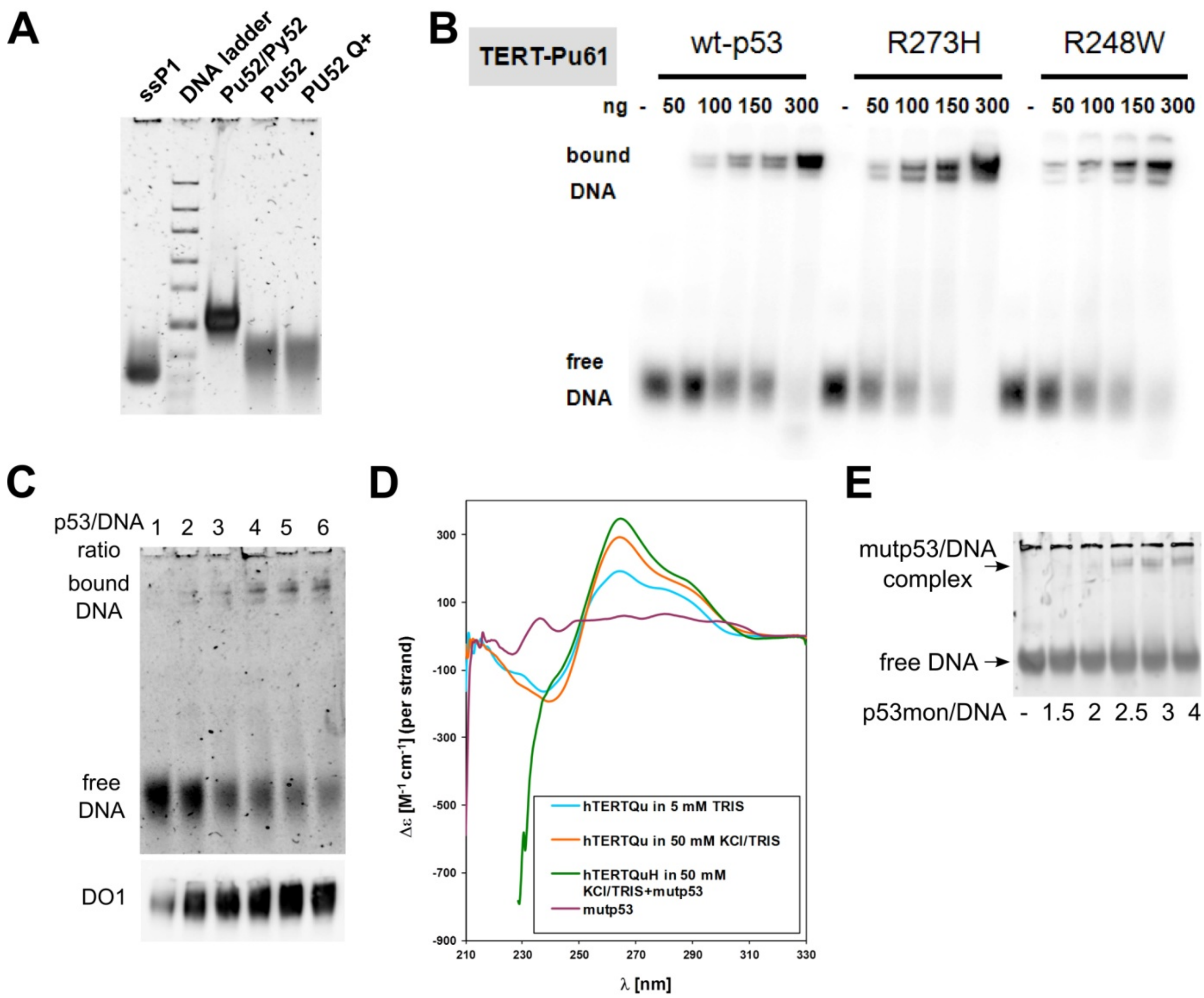

Figure S2. Analysis of p53 mutants and wtp53 interaction with quadruplex structures of the *MYC* and *TERT* promoters by EMSA and CD spectroscopy. A. Electrophoretic mobility of single-stranded P53CON-F (ssP1), dsPy52/Pu52, ssPu52 (Pu52) and ssPu52 in 50 mM KCl (Pu52Q+) (6% PA gel, 0.5xTBE). B. Comparison of the binding affinity of mutp53 and wtp53 proteins for the *TERT*-Pu61 quadruplex structure. Increasing amounts of wtp53, R248W and R273H mutp53 proteins (50-300 ng) were incubated under binding conditions (see M&M) with ^{32}P-radiolabeled *TERT*-Pu61 quadruplex structure. Samples were analyzed on a non-denaturing 6% PA gel by running in 0.5xTBE buffer supplemented with 50 mM KCl. C. EMSA control of *MYC*-Pu52 titration by mutp53R273H measured by CD spectroscopy (p53 monomer/*MYC*-Pu52 oligonucleotide ratio is 1, 2, 3, 4, 5, and 6). 10 ng of DNA mixed with protein were resolved on a 6% PA gel in 0.5xTBE, 10 mM KCl and stained with SybrGold (upper picture). Mutp53 was detected by western blotting using DO1 antibody (lower picture). D. CD spectra of *TERT*-Pu61 quadruplex structure (0.9 uM) in 5 mM Tris-HCl (blue), in 5 mM Tris-HCl supplemented with 50 mM KCl (orange), in 5 mM Tris-HCl supplemented with 219 µg mutp53R273H (3.6 µM) and 50 mM KCl (green). CD spectra were recorded 24 hours after KCl or protein addition, i.e. when equilibrium was attained. The dark violet lane represents the CD spectrum of mutp53 alone (0.5 uM). E. EMSA control of *TERT*-Pu61 titration by mutp53R273H measured by CD (p53 monomer/*TERT*-Pu61 oligonucleotide ratio is 1.5, 2, 2.5, 3, and 4). 20ng of DNA mixed with protein were resolved on 6% PA gel in 0.5xTBE, 10 mM KCl and stained with SybrGold.

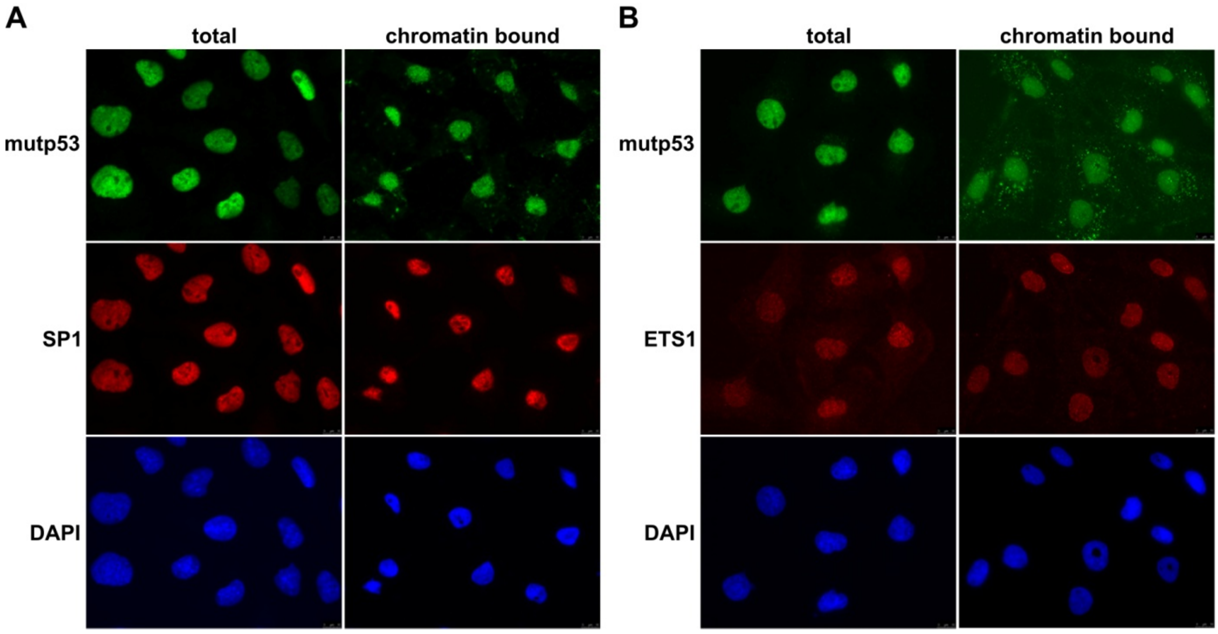

Figure S3. A, B. Immunofluorescence analysis of mutp53 co-localization with either SP1 (A) or ETS1 (B). U251 cells were grown on coverslips and either fixed directly with 4 % paraformaldehyde (left panels) or subjected to a nuclear extraction procedure prior to fixation to analyze the chromatin-associated fraction (right panels). Co-staining was performed using a p53 antibody (green) and SP1 or ETS1 antibodies (red). Nuclei were stained with DAPI (blue). All three proteins localize to the nucleus and are stably associated with chromatin.

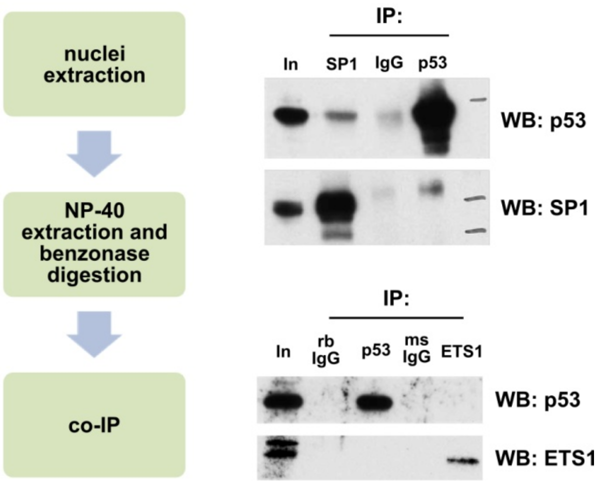

Figure S4. Co-IP analysis of mutp53, SP1 and ETS1. U251 cells were grown to ~90% confluency, harvested and nuclei were prepared to obtain the nuclear fraction for IP. Nuclei were lysed in the presence of Benzonase to extract soluble and DNA-bound proteins and IPs were performed using antibodies against p53, SP1 or ETS1. Unspecific rabbit and mouse IgG served as controls. Input and IP samples were separated by SDS-PAGE and p53, SP1 and ETS1 were detected in Western Blot.

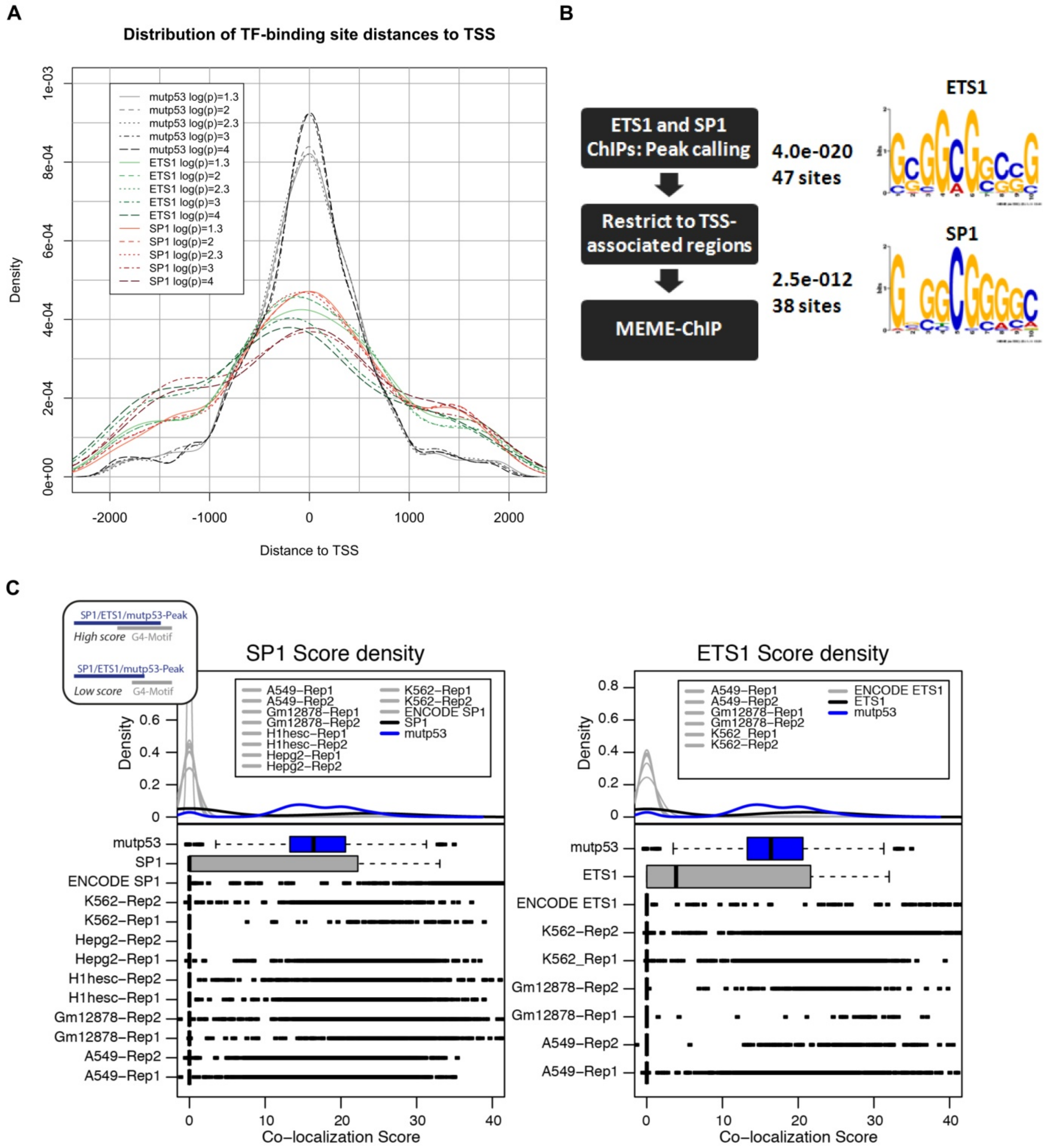

Figure S5. A. Distribution of mutp53, ETS1 and SP1 binding regions relative to TSS in dependence of a selected p-value cutoff during peak detection procedure. B. Motifs overrepresented in TSS-associated ETS1 and SP1 binding regions identified by MEME-ChIP analysis. C. Distribution of the z-scores calculated for the overlap of mutp53, ETS1 and SP1 binding regions with G4-motifs. A co-localization z-score was calculated for each binding peak rating the grade of overlap between the binding peaks and G4-motif regions in respect to either size. An increasing score value denotes a higher grade of overlap. Binding regions for mutp53, ETS1 and SP1 derived from our tiling array were compared with the ETS1 and SP1 binding regions extracted from the ENCODE database. ETS1 and SP1 ChIP-Seq data were retrieved from the ENCODE TF Binding Supertrack (contains computed peaks from different cell lines) and ENCODE/HAIB track (contains peaks from individual cell lines and replicates).

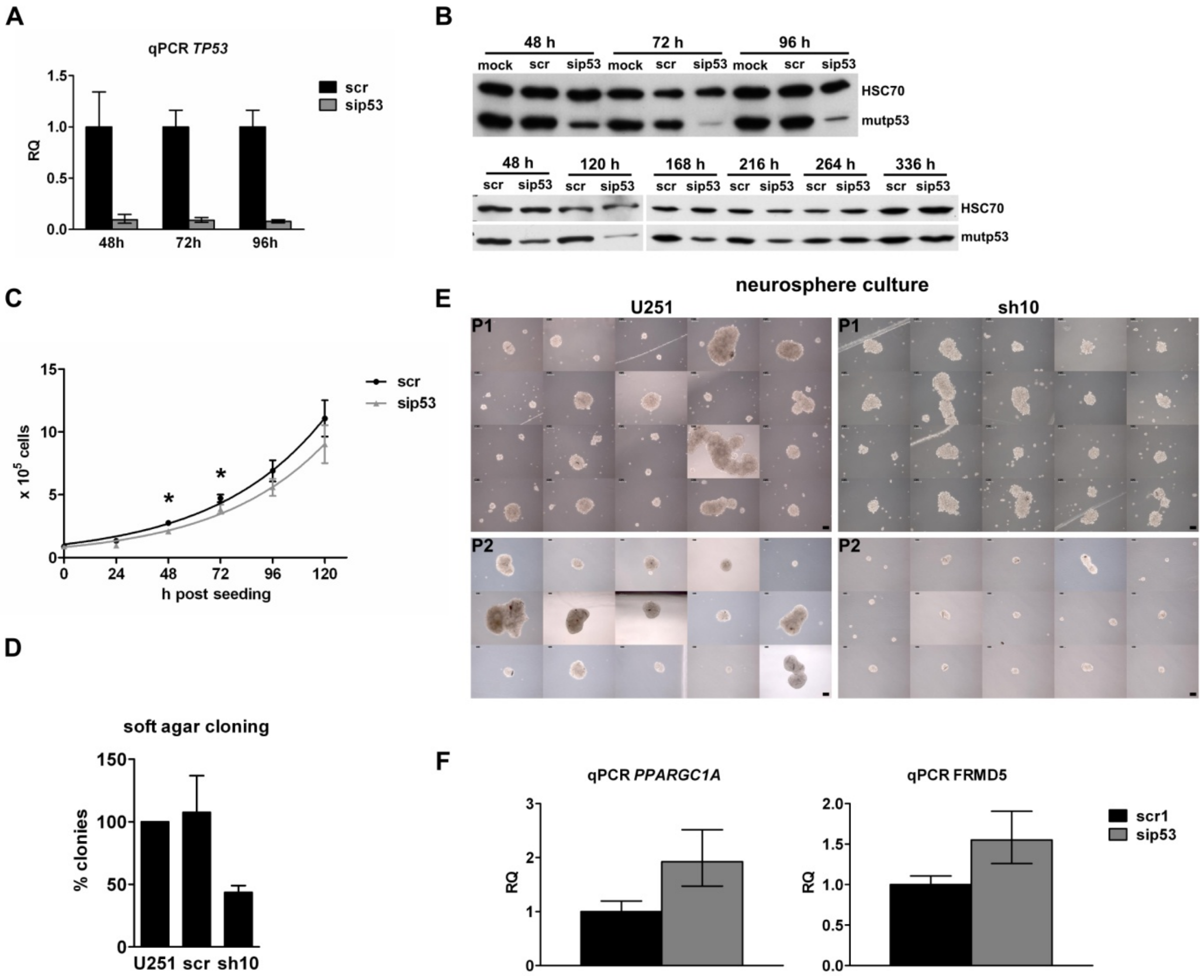

Figure S6. A. Detection of mutp53 transcripts by qPCR 48 h, 72 h and 96 h after transient transfection of p53- and scr-control-siRNA. B. WB analysis of mutp53 after transient transfection of U251 cells with p53-siRNA. Mutp53 can be efficiently depleted by RNAi. Protein levels start to diminish 48 h after transfection, show a minimum from ~72 h to ~120 h post transfection and expression levels are back to normal levels ~216 h post transfection. HSC70 was used as a loading control. C. Growth curve of p53- and scr-siRNA transfected U251 cells. Experiment was repeated in triplicates. Stars indicate p-values < 0.05. D. Soft agar colony formation of the sh10 U251 cell line (stably transfected with p53-siRNA) in comparison to parental U251 cells and a scr control cell line. After seeding in soft agar colonies in two replicate experiments were counted and are displayed as percentage of U251 colonies. E. Neurosphere formation of the sh10 U251 cell line compared to parental U251 cells in the first (P1) and second (P2) passage after seeding into non-adherent cell culture flasks. F. qPCR analysis of *PPARGC1A* and *FRMD5* gene expression after siRNA mediated mutp53 depletion in U251 cells. Transcript levels of *PPARGC1A* and *FRMD5* after mutp53 depletion relative to the scr-control (scr1) for a representative experiment are displayed.

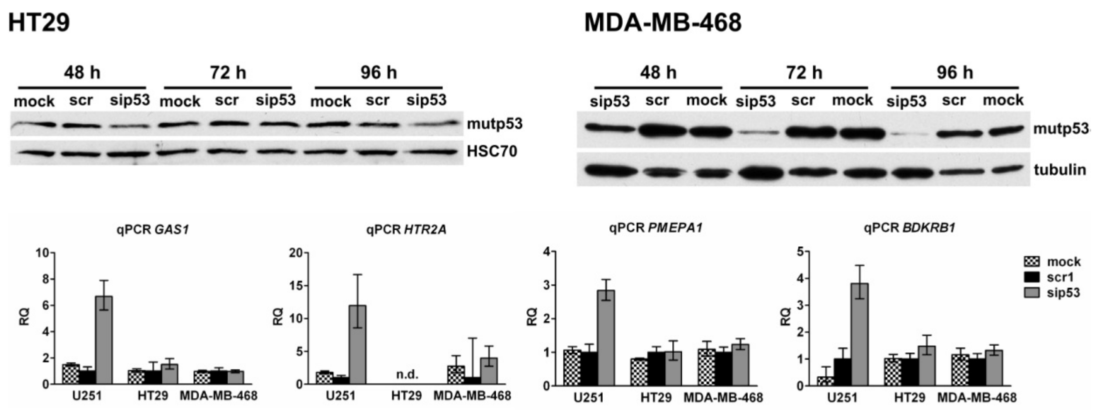

Figure S7. Analysis of *GAS1*, *HTR2A*, *BDKRB1* and *PMEPA1* gene expression after mutp53-depletion in HT29 and MDA-MB-468 cells. HT29 and MDA-MB-468 cells were transfected with p53- and scr-siRNA as well as transfection reagent alone (mock) and mutp53 protein was detected by WB analysis 48 h, 72 h and 96 h post transfection. Transcript levels were measured relative to scr-control 48 h (HT29) or 96 h (MDA-MB-468) after transfection by qPCR. Regulation of *GAS1*, *HTR2A*, *BDKRB1* and *PMEPA1* in U251 cells after mutp53 depletion is shown for comparison.

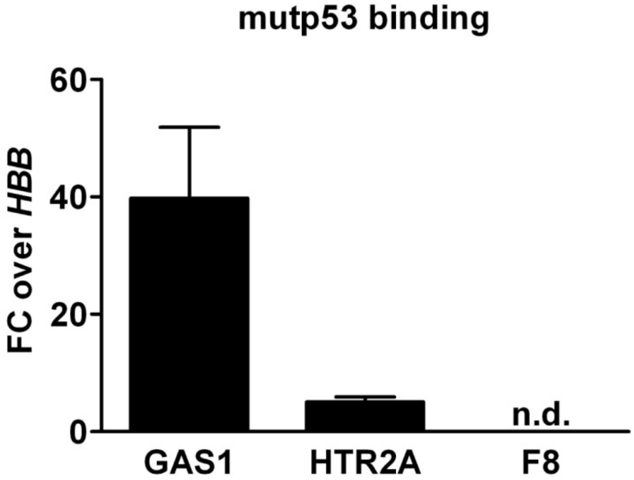

Figure S8. Validation of mutp53 binding in *GAS1* and *HTR2A* genes. DNA from three mutp53-ChIP experiments was analyzed using primers in the *GAS1* and *HTR2A* promoter regions. The tissue specific genes *F8* and *HBB* were used as background controls. Results were calculated as fold change over *HBB* and are displayed as mean +/-SEM.

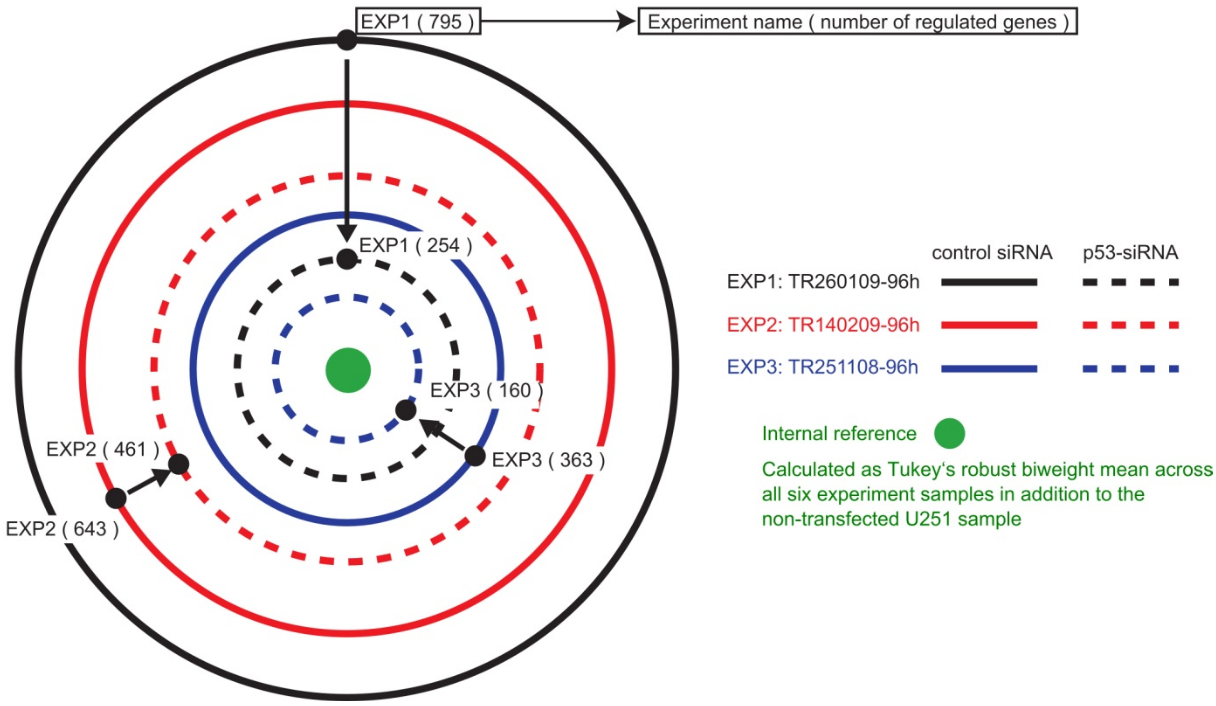

Figure S9. Model illustrating the hypothesis of an epigenetic regulation of transcription by mutant p53. The distance of the p53-specific siRNA samples to their siRNA-treated controls displays the number of transcripts regulated between the paired samples. The distance of all samples to the center reflects the number of transcripts regulated in comparison to an internal reference. The internal reference was calculated as a Tukey's biweight mean value for each transcript over the 6 siRNA-transfected samples (3x p53-specific siRNA and 3x control siRNA) and the non-transfected U251 sample. As illustrated, cells with reduced mutp53 expression exhibit a much higher similarity to each other and to the internal reference. Cells with unperturbed mutp53 expression show a higher deviation from each other and from the internal reference. As the overlap between the lists of genes regulated in the control siRNA treated samples is moderate, our model suggests a regulation similar to a rolling coin in a cone where the activity of mutp53 is capable of introducing transcriptional heterogeneity (coin circles at the top of the cone) while knockdown of mutp53 leads to a spiral like convergence (coin moves to the bottom of the cone).

SUPPLEMENTAL TABLES LEGENDS

Table S1. List of the genes, transcripts and non-coding regions represented by the Nimblegen custom 135k tiling array. The custom tiling array was designed to cover 1894 genomic intervals (Excel sheet 1) that comprise 902 coding and non-coding genes/regions (Excel sheet 2) and 1040 non-coding regions in a length of 1 kb (for use in another p53-related project). 812 tiling intervals cover single genes, 36 tiling intervals cover two genes, and 6 tiling intervals cover three genes.

Table S2. Peak calling and annotation of mutp53, SP1 and ETS1 binding regions. Binding peaks were calculated using five different p-value cutoffs (log(p-value) = 1.3, 2, 2.3, 3 and 4). Overlap of protein binding peaks with the coordinates of CpG-islands, DNaseI-HS regions, G4-motifs and Ensembl transcript start sites (TSS) was calculated as described in Methods. Description of columns for the five Excel sheets:

1) Binding Peak Name: names of the binding peaks determined using five different p-value cutoffs. ChIP-chip data for mutp53, SP1 and ETS1 were analyzed.
2) CHR Start End: chromosomal start and end positions of the binding peak
3) MaxPvalue: max. p-value as determined by CARPET peak calling procedure [1].
4) RawScore: raw score values as determined by CARPET peak calling procedure.
5) CpG-Name: name of CpG island in human genome (hg19). This information was extracted from UCSC Genome Browser CpG Islands Track. Multiple CpG islands are separated by double slash.
6) CpG-Score: intersection score between protein binding peak and CpG island. NA means no overlap between protein binding peak and CpG island.
7) CpG-Overlap: length of intersection between protein binding peak and CpG island in nucleotides.
8) DNaseI-HS region: name of Digital DNaseI Hypersensitivity (DNaseI-HS) Clusters from ENCODE (human genome, hg19). This information was extracted from UCSC Genome Browser DNase Clusters Track. Multiple regions are separated by double slash.
9) DNaseI-HS-Score: intersection score between protein binding peak and DNaseI-HS Cluster. NA means no overlap between protein binding peak and DNaseI-HS Cluster.
10) DNaseI-HS-Overlap: length of intersection between protein binding peak and DNaseI-HS Cluster in nucleotides.
11) G4-Motif: nucleotide structure of G Quadruplex Motif as calculated by non-B DB [2] (http://nonb.abcc.ncifcrf.gov/), a database that contains genomic coordinates for predicted non-B DNA forming repeats. Multiple motifs are separated by double slash. "+" and "-" signs indicate strand orientation of the G Quadruplex Motif.
12) G4-Overlap: length of intersection between protein binding peak and G Quadruplex Motif in nucleotides.
13) Tiled_region: name of tiled region from the custom tiling array (see Table S1).
14) GRCh37.p3-TSS plusStrand: TSSs of genes located on the plus strand. TSSs were named as follows: 'Associated Gene Name; Ensembl Gene ID; Transcript Biotype'. The Ensembl transcript coordinates (human genome assembly GRCh37.p3) were retrieved from Biomart

database. Multiple TSSs, which represent alternative TSS of a single gene, are separated by double slash.

15) TSS-Distance plus: distance in nucleotides between the center of the protein binding peak and the TSS of a gene located on the plus strand. To obtain this value, the genomic coordinate of the TSS was subtracted from the center position of the protein binding peak. '+' sign indicates TSS location downstream of the peak center, '-' sign indicates TSS location upstream of the peak center.
16) TSS-Distance plus medians: median distance in nucleotides between the center of the protein binding peak and alternative TSSs of a gene located on the plus strand.
17) GRCh37.p3-TSS minusStrand: TSSs of genes located on the minus strand. See description for the 'GRCh37.p3-TSS plusStrand' column.
18) TSS-Distance minus: distance in nucleotides between the center of the protein binding peak and the TSS of a gene located on the minus strand. To obtain this value, the genomic coordinate of the TSS was subtracted from the center position of the protein binding peak. '+' sign indicates TSS locationdownstream of the peak center, '-' sign indicates TSS location upstream of the peak center.
19) TSS-Distance minus medians: median distance in nucleotides between the center of the protein binding peak and alternative TSSs of a gene located on the minus strand.

Table S3. Compound list of genes differentially expressed in U251 cells which were transfected with control siRNA (scr) and p53-siRNA (sip53) in three biological replicates (TR140209, TR251108, TR260109). Description of columns in the 'compound list' Excel sheet:

1) UID Agilent probe ID
2) GeneName Name of the corresponding Ensembl gene (see the second Excel sheet for full probe annotation)
3) is regulated ('experiment name'; Yes=1, No=0): is the gene regulated 96 hours after transfection with p53-siRNA compared to cells transfected with control siRNA?
4) UP/DOWN/No regulation ('experiment name'): direction of regulation. Genes were assumed regulated if the expression value of either of the two samples compared exceeded a value of log2(100) and the absolute signal-log-ratio (SLR) was above 0.8
5) 'experiment name'-scr and 'experiment name'-sip53: name of the respective experiment (TR140209, TR251108, TR260109)
6) IsValue>log2(100) (Yes=1 or No=0): is the expression value higher than log2(100)?
7) SLR: signal-log-ratio for the respective experiment

The second Excel sheet includes re-annotation information for the Agilent probes from the 'compound list'. Re-annotation of the probes from the Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (GPL6480; last updated on July 02, 2009) was done using the GALAXY tools. The need for re-annotation of the Agilent probes arose from the observed discordance between the probe annotation provided by the manufacturer and the current annotation of genes and transcripts in the human genome. The re-annotated probes were named as follows: „S-valid" - probes that correspond to sense strands of Ensembl genes, "AS-valid" - probes that correspond to antisense strands of

Ensembl genes, “NC-valid-S” and “NC-valid-AS”- sense and antisense-specific probes associated with the non-coding (NC) fraction of the annotated Refseq transcripts.

Table S4. Peak calling and annotation of H3K4me3, H3K9ac, RNA Pol II S2-P and and RNA Pol II S5-P enriched regions. Parameters used during CARPET peak calling procedure are shown in sheet 1. Overlap of enriched regions with the coordinates of CpG-islands, DNaseI-HS regions, G4-motifs, Ensembl transcript start sites (TSS) (sheets 2 and 3) was calculated as described in Methods. ChIP-chip analysis was performed using U251 cells transfected with control siRNA (scr) and p53-siRNA (sip53). For description of columns in the Excel sheets 2 and 3 see the legend to Table S2.

Table S5. List of the reagents (primers, oligonucleotides, siRNA and antibodies) used in the study.

Table S6. Intersection between mutp53-binding regions and histone modifications (H3K4me3, H3K9ac) and RNA Pol II S2-P/S5-P binding peaks. This table includes only mutp53-binding regions associated with TSSs (see Supplementary Materials and Methods for details). ChIP-chip analysis was performed using U251 cells transfected with control siRNA (scr) and p53-siRNA (sip53). For description of columns see the legend to Table S2. Additional columns include calculated ratios (from 0 to 1; NA – no overlap) of overlaps between mutp53-binding regions and histone modifications and RNA Pol II S2-P/S5-P binding peaks.

REFERENCES

1. Cesaroni M, Cittaro D, Brozzi A, Pelicci PG, Luzi L. CARPET: a web-based package for the analysis of ChIP-chip and expression tiling data. Bioinformatics 2008; 24:2918-20.
2. Cer RZ, Bruce KH, Mudunuri US, Yi M, Volfovsky N, Luke BT, et al. Non-B DB: a database of predicted non-B DNA-forming motifs in mammalian genomes. Nucleic Acids Res 2011; 39:D383-91.
3. Myers RM, (ENCODE-Project-Consortium). A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol 2011; 9:e1001046.
4. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, et al. Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol 2010; Chapter 19:Unit 19 0 1-21.
5. Goecks J, Nekrutenko A, Taylor J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol 2010; 11:R86.
6. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol 2011; 29:24-6.
7. Bailey TL, Boden M, Whitington T, Machanick P. The value of position-specific priors in motif discovery using MEME. BMC Bioinformatics 2010; 11:179.
8. Bock C, Halachev K, Buch J, Lengauer T. EpiGRAPH: user-friendly software for statistical analysis and prediction of (epi)genomic data. Genome Biol 2009; 10:R14.
9. Renciuk D, Kejnovska I, Skolakova P, Bednarova K, Motlova J, Vorlickova M. Arrangements of human telomere DNA quadruplex in physiologically relevant K+ solutions. Nucleic Acids Res 2009; 37:6625-34.
10. Miroux B, Walker JE. Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 1996; 260:289-98.
11. Brazdova M, Kizek R, Havran L, Palecek E. Determination of glutathione-S-transferase traces in preparations of p53 C-terminal domain (aa320-393). Bioelectrochemistry 2002; 55:115-8.
12. Brazdova M, Palecek J, Cherny DI, Billova S, Fojta M, Pecinka P, et al. Role of tumor suppressor p53 domains in selective binding to supercoiled DNA. Nucleic Acids Res 2002; 30:4966-74.
13. Gray DM, Hung SH, Johnson KH. Absorption and circular dichroism spectroscopy of nucleic acid duplexes and triplexes. Methods Enzymol 1995; 246:19-34.